

# Pemodelan Faktor-Faktor yang Memengaruhi Indeks Pembangunan Kesehatan Masyarakat Provinsi Jawa Timur Menggunakan Pendekatan Regresi Semiparametrik *Spline*

Made Ayu Dwi Octavanny, I Nyoman Budiantara, dan Vita Ratnasari

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: i\_nyoman\_b@statistika.its.ac.id, vita\_ratna@statistika.its.ac.id

**Abstrak**—Aspek kesehatan yang dilihat dari umur harapan hidup saat kelahiran belum dapat menjabarkan indeks kesehatan dengan baik sehingga diperlukan indikator yang dapat mengukur aspek kesehatan, yaitu Indeks Pembangunan Kesehatan Masyarakat (IPKM). IPKM dapat digunakan untuk mengetahui sejauh mana keberhasilan pembangunan kesehatan masyarakat suatu wilayah. Hampir semua kabupaten/kota di Provinsi Jawa Timur mengalami penurunan peringkat berdasarkan pengembangan IPKM 2013. Hal ini perlu menjadi perhatian bagi pemerintah untuk mengetahui faktor-faktor yang memengaruhi IPKM di Provinsi Jawa Timur sehingga pemerintah dapat melakukan usaha perbaikan kualitas kesehatan masyarakat. Pada penelitian ini, IPKM dan faktor-faktor yang memengaruhinya sebagian membentuk pola tertentu dan sebagian lagi tidak membentuk pola tertentu, sehingga digunakan metode regresi semiparametrik *spline*. Model terbaik didapatkan dari titik knot optimum berdasarkan nilai *Generalized Cross Validation* (GCV) minimum. Model regresi semiparametrik *spline* terbaik adalah dengan menggunakan tiga titik knot. Terdapat empat variabel yang signifikan, yaitu Angka Kematian Bayi, kepadatan penduduk, persentase rumah tangga berperilaku hidup bersih dan sehat, dan persentase rumah sehat. Nilai koefisien determinasi yang dihasilkan dari model ini adalah sebesar 92,99%.

**Kata Kunci**—Indeks Pembangunan Kesehatan Masyarakat, Jawa Timur, Regresi Semiparametrik *Spline*, *Spline Truncated*.

## I. PENDAHULUAN

**T**ERCAPAINYA pembangunan sumber daya manusia ditandai dengan tingginya kualitas hidup yang dicapai masyarakat. Salah satu standar yang ditetapkan untuk mengukur sejauh mana keberhasilan pembangunan manusia adalah Indeks Pembangunan Manusia (IPM). IPM dibentuk melalui pendekatan tiga dimensi dasar, yaitu lamanya hidup, pengetahuan, dan penghidupan yang layak [1]. Aspek kesehatan yang dilihat dari umur harapan hidup saat kelahiran belum dapat menjabarkan indeks kesehatan dengan baik sehingga diperlukan indikator yang dapat mengukur pembangunan kesehatan, yaitu Indeks Pembangunan Kesehatan Masyarakat (IPKM). IPKM merupakan indikator komposit yang menggambarkan kemajuan pembangunan kesehatan yang dirumuskan dari 30 indikator kesehatan. Indeks tersebut disusun oleh Badan Penelitian dan

Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan RI.

Indeks kesehatan Indonesia menduduki peringkat ke- 108 dari seluruh negara di dunia menurut *Human Development Reports* pada tahun 2013. Hampir semua kabupaten/kota di Provinsi Jawa Timur mengalami penurunan peringkat berdasarkan pengembangan IPKM 2013 [2]. Oleh karena itu, perlu diketahui faktor-faktor yang memengaruhi IPKM di Provinsi Jawa Timur sehingga pemerintah dapat melakukan upaya-upaya untuk meningkatkan IPKM sebagai usaha perbaikan kualitas kesehatan masyarakat.

Salah satu metode yang dapat digunakan untuk mengetahui pola hubungan antara variabel respon dengan prediktor dimana sebagian mengikuti pola tertentu dan sebagian lagi tidak mengikuti pola tertentu adalah regresi semiparametrik *spline* [3]. Penelitian mengenai indeks kesehatan di Provinsi Jawa Timur sebelumnya telah dilakukan oleh Riskiyanti [4] menggunakan regresi multivariat dan Maully [5] dengan regresi logistik biner. Selain itu, Prasetyo [6] melakukan penelitian mengenai data kesehatan Kabupaten Banyuwangi dengan regresi terboboti geografis, serta Diansuantari [7] meneliti mengenai derajat kesehatan masyarakat Provinsi Bali dengan metode MARS.

## II. TINJAUAN PUSTAKA

### A. Pengujian Nonlinieritas

Untuk mengetahui apakah dua variabel memiliki hubungan linier atau nonlinier, maka dapat dilakukan dengan pengujian nonlinieritas. Salah satu pengujian nonlinieritas adalah uji Ramsey's RESET [8]. Pengujian ini dilakukan untuk membantu pembuatan keputusan apakah kedua variabel memiliki hubungan linier atau nonlinier yang seringkali belum dapat diputuskan jika hanya berdasarkan *scatterplot*.

### B. Regresi Semiparametrik

Analisis regresi adalah suatu metode Statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan prediktor [9]. Terdapat tiga pendekatan dalam analisis regresi, yaitu pendekatan regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Budiantara [10] mendefinisikan regresi semiparametrik sebagai model regresi

yang bentuk kurva regresi sebagian diasumsikan diketahui polanya dan sebagian tidak diketahui bentuk polanya. Misal terdapat data berpasangan  $(x_i, y_i, t_i)$  dan hubungan antara data berpasangan tersebut mengikuti regresi semiparametrik. Berikut merupakan model regresi semiparametrik [3].

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + f(t_i) + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (1)$$

Variabel respon  $y_i$  berhubungan parametrik dengan variabel prediktor  $x_i$  dan berhubungan nonparametrik dengan variabel prediktor  $t_i$ . Error  $\varepsilon_i$  diasumsikan identik, independen, dan berdistribusi normal.

### C. Pemodelan Regresi Spline Truncated

Model *spline* memiliki fleksibilitas yang tinggi dan cenderung mencari sendiri estimasi data ke mana pola data tersebut bergerak. Kelebihan ini terjadi karena dalam *spline* terdapat titik-titik knot. Misalkan  $y_i$  merupakan variabel respon,  $i=1, 2, \dots, n$  sedangkan  $x_{1i}, x_{2i}, \dots, x_{pi}$  adalah variabel prediktor yang tidak diketahui pola hubungannya dengan variabel respon. Hubungan  $x_{1i}, x_{2i}, \dots, x_{pi}, y_i$  dapat dimodelkan dengan regresi nonparametrik multivariabel sebagai berikut.

$$y_i = f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_p(x_{pi}) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2)$$

Pada regresi *spline*, masing-masing fungsi  $f$  dapat ditulis sebagai berikut.

$$f(x_{li}) = \sum_{j=0}^p \beta_{lj} x_{li}^j + \sum_{k=1}^K \beta_{l(p+k)} (x_{li} - \lambda_{lk})_+^p \quad (3)$$

Fungsi potongan (*truncated*) sebagai berikut.

$$(x_{li} - \lambda_{lk})_+^p = \begin{cases} (x_{li} - \lambda_{lk})^p, & x_{li} \geq \lambda_{lk} \\ 0, & x_{li} < \lambda_{lk} \end{cases} \quad (4)$$

Error  $\varepsilon_i$  diasumsikan identik, independen, dan berdistribusi normal.

### D. Pemilihan Titik Knot Optimal

Titik knot adalah titik perpaduan bersama dimana terdapat perubahan perilaku pola data pada interval yang berlainan [11]. Salah satu metode yang dapat digunakan untuk menentukan titik knot optimal adalah *Generalized Cross Validation* (GCV). Titik knot optimal diperoleh dari nilai GCV minimum. Secara umum, fungsi GCV adalah sebagai berikut.

$$GCV(\lambda_1, \lambda_2, \dots, \lambda_K) = \frac{MSE(\lambda_1, \lambda_2, \dots, \lambda_K)}{(n^{-1} \text{trace}[I - A(\lambda_1, \lambda_2, \dots, \lambda_K)])^2} \quad (5)$$

dimana  $\mathbf{I}$  adalah matriks identitas,  $n$  jumlah pengamatan,  $A(\lambda_1, \lambda_2, \dots, \lambda_K)$  merupakan matriks  $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ , serta  $MSE(\lambda_1, \lambda_2, \dots, \lambda_K)$  diberikan oleh:

$$MSE(\lambda_1, \lambda_2, \dots, \lambda_K) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

### E. Pengujian Parameter Model

Pengujian parameter model digunakan untuk mengetahui apakah variabel prediktor berpengaruh atau tidak terhadap variabel respon. Diberikan model regresi semiparametrik sebagai berikut.

$$y_i = \sum_{j=0}^q \beta_j x_{ji} + \sum_{l=1}^p \gamma_l t_i^l + \sum_{k=1}^r \gamma_{p+k} (t_i - \lambda_k)_+^p + \varepsilon_i \quad (7)$$

### 1. Pengujian Secara Serentak

Pengujian secara serentak dilakukan untuk mengetahui apakah variabel prediktor secara bersama-sama berpengaruh terhadap variabel respon. Hipotesis yang digunakan dalam pengujian serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_q = \gamma_1 = \gamma_2 = \dots = \gamma_{p+r} = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_j \neq 0 \text{ atau } \gamma_l \neq 0, \quad j=1, 2, \dots, q, \quad l=1, 2, \dots, p+r$$

Statistik uji yang digunakan adalah sebagai berikut.

$$F = \frac{MS_{regresi}}{MS_{error}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / (q + p + r)}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - (q + p + r) - 1)} \quad (8)$$

dengan daerah penolakan  $H_0$  adalah  $F > F_{(\alpha, (q+p+r), (n-(q+p+r)-1))}$  atau  $P\text{-value} < \alpha$ .

### 2. Pengujian Secara Parsial

Hipotesis yang digunakan pada pengujian parsial untuk komponen parametrik adalah sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, \quad j = 1, 2, \dots, q$$

Statistik uji yang digunakan untuk komponen parametrik adalah sebagai berikut.

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}, \quad j = 1, 2, \dots, q \quad (9)$$

Selanjutnya, hipotesis yang digunakan dalam pengujian parsial untuk komponen nonparametrik adalah sebagai berikut.

$$H_0 : \gamma_l = 0$$

$$H_1 : \gamma_l \neq 0, \quad l = 1, 2, \dots, p+r$$

Statistik uji yang digunakan untuk komponen nonparametrik adalah sebagai berikut.

$$t = \frac{\hat{\gamma}_l}{SE(\hat{\gamma}_l)}, \quad l = 1, 2, \dots, p+r \quad (10)$$

Daerah penolakan untuk komponen parametrik dan komponen nonparametrik adalah  $|t| > t_{(\alpha/2, n-(q+p+r)-1)}$  atau  $P\text{-value} < \alpha$ .

### F. Kriteria Keabakan Model

Salah satu kriteria yang dapat digunakan untuk menentukan kebaikan model regresi adalah menggunakan koefisien determinasi ( $R^2$ ) dengan rumus sebagai berikut.

$$R^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / \sum_{i=1}^n (y_i - \bar{y})^2 \quad (11)$$

### G. Pengujian Asumsi Residual

Pengujian asumsi residual dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi identik, independen, dan berdistribusi normal (IIDN).

#### 1. Uji Asumsi Identik

Pengujian asumsi identik bertujuan untuk mengetahui homogenitas variansi residual. Pengujian asumsi identik dilakukan dengan uji Glejser [12]. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, \quad i = 1, 2, \dots, n$$

Statistik uji Glejser dirumuskan sebagai berikut.

$$F_{hitung} = \frac{\sum_{i=1}^n \left( \hat{e}_i - |\bar{e}|^2 \right) / (q+p+r)}{\sum_{i=1}^n \left( |e_i| - |\hat{e}_i|^2 \right) / (n-(q+p+r)-1)} \quad (12)$$

dengan daerah penolakan adalah  $F > F_{(\alpha, (q+p+r), (n-(q+p+r)-1))}$ .

## 2. Asumsi Independen

Asumsi independen dilakukan untuk mengetahui apakah terdapat autokorelasi atau tidak pada residual. Untuk mengetahui apakah terjadi autokorelasi secara visual dilakukan dengan melihat plot dari *Autocorrelation Function* (ACF) dari residual. Untuk mendapatkan nilai ACF digunakan rumus sebagai berikut.

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^n (e_{t-k} - \bar{e})(e_t - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (13)$$

Asumsi independen dapat dideteksi dengan menggunakan interval konfidensi dengan rumus sebagai berikut.

$$-t_{\alpha/2, n-1} SE(\hat{\rho}_k) < \rho_k < t_{\alpha/2, n-1} SE(\hat{\rho}_k) \quad (14)$$

Apabila terdapat ACF yang keluar dari interval konfidensi, maka diindikasikan adanya autokorelasi antar residual.

## 3. Uji Asumsi Berdistribusi Normal

Untuk mengetahui apakah residual berdistribusi normal, dapat dilakukan dengan pengujian *Kolmogorov-Smirnov* [13]. Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : F(x) = F_0(x)$  atau residual berdistribusi normal

$H_1 : F(x) \neq F_0(x)$  atau residual tidak berdistribusi normal

Nilai statistik uji *Kolmogorov-Smirnov* adalah sebagai berikut.

$$D = \sup_x |S(x) - F_0(x)| \quad (15)$$

Daerah penolakan adalah  $|D| > q_{(1-\alpha)}$  pada tabel *Kolmogorov-Smirnov* atau  $P\text{-value} < \alpha$ .

## H. Indeks Pembangunan Kesehatan Masyarakat

IPKM adalah indikator komposit yang menggambarkan kemajuan pembangunan kesehatan yang dirumuskan dari data kesehatan, yaitu Riskesdas, Susenas, dan Podes. IPKM pertama dirumuskan dari 24 indikator kesehatan yang kemudian dilakukan pengembangan model IPKM 2013 dengan menggunakan 30 indikator. IPKM dapat dimanfaatkan untuk menentukan peringkat provinsi dan kabupaten/kota dalam keberhasilan pembangunan kesehatan masyarakat, sebagai bahan advokasi ke pemerintah sehingga sumber daya dan program kesehatan diprioritaskan, serta sebagai salah satu kriteria penentuan alokasi dana bantuan kesehatan pusat ke daerah [2].

## III. METODOLOGI PENELITIAN

### A. Sumber Data

Data yang digunakan adalah data sekunder yang diambil dari Badan Pusat Statistika Jawa Timur dan publikasi Profil Kesehatan Jawa Timur. Data yang digunakan merupakan data pada tahun 2013 dengan banyak observasi adalah 38 kabupaten/kota yang ada di Provinsi Jawa Timur.

### B. Variabel Penelitian

Berikut merupakan variabel yang digunakan pada penelitian ini.

Tabel 1.  
Variabel Penelitian

Variabel	Keterangan
y	IPKM
$x_1$	Persentase Penduduk Miskin
$x_2$	Angka Kematian Bayi
$t_1$	Kepadatan Penduduk
$t_2$	Angka Kematian Ibu
$t_3$	Persentase Rumah Tangga Ber-PHBS
$t_4$	Persentase Rumah Sehat

### C. Langkah-Langkah Penelitian

Langkah-langkah dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan data IPKM di Provinsi Jawa Timur beserta variabel-variabel yang memengaruhi IPKM.
2. Mendeskripsikan karakteristik data dari IPKM di Provinsi Jawa Timur beserta variabel-variabel yang memengaruhi IPKM.
3. Membuat *scatterplot* dan melakukan uji RESET untuk mengetahui pola data antara variabel IPKM dengan variabel-variabel prediktornya.
4. Mencari variabel prediktor komponen parametrik.
5. Mencari variabel prediktor komponen nonparametrik.
6. Memodelkan data dengan pendekatan *spline* dengan satu knot, dua knot, tiga knot, dan kombinasi knot.
7. Memilih titik knot optimal berdasarkan nilai *Generalized Cross Validation* (GCV) yang minimum.
8. Mendapatkan model regresi *spline* dengan titik knot optimal.
9. Menguji signifikansi parameter regresi *spline* secara serentak dan parsial. Menghitung nilai koefisien determinasi ( $R^2$ ).
10. Menguji asumsi residual IIDN. Apabila residual model *spline* tidak memenuhi asumsi, maka dilakukan transformasi data. Selanjutnya memulai kembali dari langkah analisis (6).
11. Menginterpretasikan model dan menarik kesimpulan.

## IV. ANALISIS DAN PEMBAHASAN

### A. Karakteristik Indeks Pembangunan Kesehatan Masyarakat di Provinsi Jawa Timur

Berikut ini adalah karakteristik dari keenam variabel yang diduga memengaruhi IPKM di Provinsi Jawa Timur yang disajikan pada Tabel 2.

Tabel 2.  
Statistika Deskriptif Variabel Penelitian

Variabel	Rata-rata	Varians	Minimum	Maksimum
y	0,70487	0,00254	0,5874	0,79
$x_1$	12,54	27,136	4,77	27,08
$x_2$	32,79	155,25	18,71	62,45
$t_1$	1802	4664624	382	8551
$t_2$	105,23	1856,33	30,82	212,71
$t_3$	45,34	210,75	17,14	67,32
$t_4$	38,85	548,91	1,02	81,03

Tabel 2 menunjukkan bahwa variabel respon (y) merupakan IPKM di Provinsi Jawa Timur dengan rata-rata IPKM pada tahun 2013 sebesar 0,70487. IPKM terendah di Provinsi Jawa Timur terdapat di Kabupaten Pamekasan dengan IPKM sebesar 0,5874. Indeks kesehatan di Kabupaten

Pamekasan rendah disebabkan oleh kurangnya pengetahuan masyarakat tentang gizi bayi dan balita sehingga kasus gizi buruk dan gizi kurang masih banyak terjadi. Selain itu, masyarakat di Kabupaten Pamekasan masih kurang sadar akan pentingnya perilaku hidup bersih dan sehat. Kota Madiun meraih IPKM tertinggi di Jawa Timur dengan nilai sebesar 0,79. Hal ini disebabkan oleh fasilitas kesehatan yang baik dari segi kualitas maupun kuantitas masyarakat di Kota Madiun telah memadai bahkan menjadi rujukan bagi rumah sakit di kabupaten sekitarnya. Selain itu angka harapan hidup Kota Madiun yang terus meningkat sejak tahun 2009 hingga tahun 2013 yang menandakan kesehatan penduduk yang semakin baik.

Variabel  $x_1$  merupakan variabel persentase penduduk miskin dengan rata-rata sebesar 12,54%. Persentase tertinggi terdapat di Kabupaten Sampang, sedangkan persentase penduduk miskin terendah terdapat di Kota Batu. Variabel  $x_2$  merupakan variabel Angka Kematian Bayi (AKB) yang memiliki nilai rata-rata sebesar 32,79. AKB tertinggi terdapat pada Kabupaten Probolinggo dan AKB terendah terdapat di Kota Blitar. Variabel  $t_1$  merupakan variabel kepadatan penduduk yang memiliki rata-rata sebesar 1802 per km<sup>2</sup>. Nilai tertinggi terdapat di Kota Surabaya dan kepadatan penduduk terendah terdapat di Kabupaten Pacitan. Variabel  $t_2$  merupakan variabel Angka Kematian Ibu (AKI) dengan nilai rata-rata sebesar 105,23. AKI tertinggi terdapat di Kota Probolinggo, sedangkan AKI terendah terdapat di Kota Batu. Variabel  $t_3$  merupakan variabel persentase rumah tangga berperilaku hidup bersih dan sehat yang memiliki rata-rata sebesar 45,34%. Persentase tertinggi terdapat di Kota Surabaya dan persentase terendah terdapat di Kabupaten Situbondo. Variabel  $t_4$  merupakan variabel persentase rumah sehat yang memiliki rata-rata sebesar 38,85%. Persentase tertinggi terdapat di Kota Surabaya, sedangkan persentase rumah sehat terendah terdapat di Kabupaten Blitar.

Terdapat 13 kabupaten/kota yang masih berada di bawah IPKM Indonesia, yaitu 0,6879. Walaupun sebagian besar Kabupaten/Kota Provinsi Jawa Timur sudah memiliki nilai yang lebih besar dari IPKM Indonesia, namun IPKM di Jawa Timur masih perlu ditingkatkan dengan memerhatikan faktor-faktor yang memengaruhi IPKM.

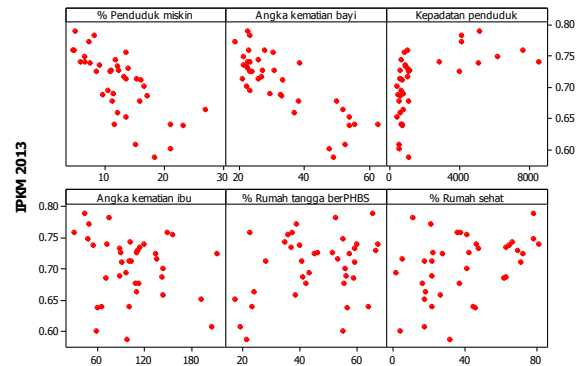
### B. Pola Hubungan IPKM di Provinsi Jawa Timur dengan Faktor yang Diduga Memengaruhi

Gambar 1 menunjukkan bahwa pola hubungan antara IPKM dengan persentase penduduk miskin dan Angka Kematian Bayi cenderung membentuk suatu pola tertentu sehingga termasuk ke dalam komponen parametrik. Sedangkan pola hubungan antara IPKM dengan kepadatan penduduk, Angka Kematian Ibu, persentase rumah tangga ber-PHBS, dan persentase rumah sehat tidak membentuk pola tertentu sehingga termasuk ke dalam komponen nonparametrik.

Selain menggunakan *scatterplot*, untuk mengetahui hubungan antara IPKM dengan faktor-faktor yang memengaruhinya dapat menggunakan uji RESET.

Berdasarkan Tabel 3, didapatkan kesimpulan yang berbeda dengan *scatterplot* yang telah dilakukan sebelumnya. Dengan menggunakan uji RESET, variabel yang termasuk

dalam komponen parametrik adalah persentase penduduk miskin, Angka Kematian Bayi, Angka Kematian Ibu, dan persentase rumah sehat, sedangkan variabel yang termasuk dalam komponen nonparametrik adalah kepadatan penduduk dan persentase rumah tangga ber-PHBS. Berdasarkan hal tersebut, metode yang digunakan menggunakan pendekatan regresi semiparametrik *spline* karena terdapat komponen parametrik dan komponen nonparametrik.



Gambar 1. Pola Hubungan IPKM dengan Variabel Prediktor

Tabel 3.  
Uji RESET Antara IPKM dengan Faktor yang Memengaruhi

Variabel	F	P-value	Kesimpulan
$x_1$	1,8377	0,1839	Gagal tolak, linier
$x_2$	0,77701	0,3841	Gagal tolak, linier
$t_1$	4,7216	0,03664	Tolak, nonlinier
$t_2$	0,055805	0,8146	Gagal tolak, linier
$t_3$	3,004	0,09186	Tolak, nonlinier
$t_4$	0,38616	0,5384	Gagal tolak, linier

### C. Pemilihan Titik Knot Optimum

Berikut akan ditampilkan pada Tabel 4 nilai GCV minimum dengan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi titik knot baik untuk pola berdasarkan *scatterplot* dan pola berdasarkan uji RESET.

Tabel 4.  
Perbandingan Nilai GCV Minimum

Keterangan	Jumlah Knot	GCV Minimum
Pola Berdasarkan Scatterplot	Satu Titik Knot	0,0008940
	Dua Titik Knot	0,0007568
	Tiga Titik Knot	0,0005271
	Kombinasi Knot (1,3,3,3)	0,0005568
Pola Berdasarkan Uji RESET	Satu Titik Knot	0,000837
	Dua Titik Knot	0,0007515
	Tiga Titik Knot	0,0005639
	Kombinasi Knot (1,3)	0,0005278

Berdasarkan Tabel 4, diketahui bahwa pemodelan yang menghasilkan nilai GCV paling minimum adalah pada pola berdasarkan *scatterplot* dengan menggunakan tiga titik knot untuk masing-masing variabel komponen nonparametrik.

### D. Pengujian Signifikansi Parameter Model Regresi Semiparametrik Spline

#### 1. Pengujian Serentak

Pengujian secara serentak dilakukan pada parameter model regresi terhadap variabel IPKM secara bersama-sama atau serentak. Hasil pengujian secara serentak ditampilkan pada Tabel 5.

Tabel 5.  
Hasil Pengujian Serentak

Sumber	df	SS	MS	F	P-value
Regresi	18	0,0884	0,005	16,81	0,00
Error	19	0,0056	0,0003		
Total	37	0,0940			

Berdasarkan Tabel 5, diketahui bahwa diperoleh nilai F sebesar 16,81 dan *P-value* sebesar 0. Dengan taraf signifikansi ( $\alpha$ ) sebesar 5% maka didapatkan keputusan tolak  $H_0$  sehingga dapat disimpulkan bahwa minimal terdapat satu parameter pada model yang signifikan. Nilai  $R^2$  yang diperoleh adalah 94,09%.

## 2. Pengujian Individu

Berikut merupakan hasil dari pengujian parameter model regresi secara individu yang disajikan pada Tabel 6. Berdasarkan Tabel 6, diketahui bahwa dengan taraf signifikansi ( $\alpha$ ) sebesar 5%, terdapat empat variabel yang signifikan dan dua variabel yang tidak signifikan. Empat variabel yang signifikan adalah variabel Angka Kematian Bayi ( $x_2$ ), kepadatan penduduk ( $t_1$ ), persentase rumah tangga ber-PHBS ( $t_3$ ), dan persentase rumah sehat ( $t_4$ ).

Tabel 6.  
Hasil Pengujian Individu

Variabel	Parameter	t	Keputusan	Kesimpulan
Konstan	$\beta_0$	12,699	Tolak	Signifikan
$x_1$	$\beta_1$	-0,863	Gagal	Tidak Signifikan
$x_2$	$\beta_2$	-6,883	Tolak	Signifikan
$t_1$	$\gamma_1$	1,695	Gagal	Signifikan
	$\gamma_2$	1,418	Gagal	
	$\gamma_3$	-2,580	Tolak	
	$\gamma_4$	-1,899	Gagal	
$t_2$	$\gamma_5$	1,393	Gagal	Tidak Signifikan
	$\gamma_6$	-0,935	Gagal	
	$\gamma_7$	1,147	Gagal	
	$\gamma_8$	-1,375	Gagal	
$t_3$	$\gamma_9$	-2,690	Tolak	Signifikan
	$\gamma_{10}$	3,010	Tolak	
	$\gamma_{11}$	-3,415	Tolak	
	$\gamma_{12}$	3,349	Tolak	
$t_4$	$\gamma_{13}$	0,313	Gagal	Signifikan
	$\gamma_{14}$	-3,151	Tolak	
	$\gamma_{15}$	3,360	Tolak	
	$\gamma_{16}$	-3,414	Tolak	

## E. Pemilihan Titik Knot Optimum dengan Empat Variabel

Selanjutnya melakukan pemodelan kembali dengan menghapus variabel yang tidak signifikan. Berikut akan ditampilkan pada Tabel 7 nilai GCV minimum pada model dengan empat variabel.

Tabel 7.

Perbandingan Nilai GCV Minimum	
Jumlah Knot	GCV Minimum
Satu Titik Knot	0,0008246
Dua Titik Knot	0,0006411
Tiga Titik Knot	0,0004339
Kombinasi Knot (3,1,3)	0,0004438

Berdasarkan Tabel 7, diketahui bahwa pemodelan yang menghasilkan nilai GCV paling minimum adalah dengan menggunakan tiga titik knot untuk masing-masing variabel komponen nonparametrik.

## F. Pengujian Signifikansi Parameter Model Regresi Semiparametrik Spline

### 1. Pengujian Serentak

Pengujian secara serentak dilakukan pada parameter model regresi terhadap variabel IPKM secara bersama-sama atau serentak. Hasil pengujian secara serentak ditampilkan pada Tabel 8.

Tabel 8. Hasil Pengujian Serentak Pada Model Empat Variabel					
Sumber	df	SS	MS	F	P-value
Regresi	13	0,087318	0,006717	24,51	0,00
Error	24	0,006577	0,000274		
Total	37	0,093895			

Berdasarkan Tabel 8, diketahui bahwa diperoleh nilai F sebesar 24,51 dan *P-value* sebesar 0. Dengan taraf signifikansi ( $\alpha$ ) sebesar 5% maka didapatkan keputusan tolak  $H_0$  sehingga dapat disimpulkan bahwa minimal terdapat satu parameter pada model yang signifikan. Nilai  $R^2$  yang diperoleh adalah 92,99%.

### 2. Pengujian Individu

Berikut merupakan hasil dari pengujian parameter model regresi secara individu yang disajikan pada Tabel 9.

Tabel 9. Hasil Pengujian Individu Pada Model Empat Variabel				
Variabel	Parameter	t	Keputusan	Kesimpulan
Konstan	$\beta_0$	13,8592	Tolak	Signifikan
$x_2$	$\beta_2$	-8,4150	Tolak	Signifikan
$t_1$	$\gamma_1$	3,9135	Tolak	Signifikan
	$\gamma_2$	-3,0295	Tolak	
	$\gamma_3$	3,0541	Tolak	
	$\gamma_4$	-3,1215	Tolak	
$t_3$	$\gamma_5$	-0,2756	Gagal	Signifikan
	$\gamma_6$	1,0270	Gagal	
	$\gamma_7$	-1,9891	Gagal	
	$\gamma_8$	2,2049	Tolak	
$t_3$	$\gamma_9$	1,6401	Gagal	Signifikan
	$\gamma_{10}$	-5,7326	Tolak	
	$\gamma_{11}$	6,3390	Tolak	
	$\gamma_{12}$	-6,2259	Tolak	

Berdasarkan Tabel 9, diketahui bahwa dengan taraf signifikansi ( $\alpha$ ) sebesar 5%, semua variabel prediktor berpengaruh signifikan terhadap model.

## G. Pemodelan IPKM dengan Titik Knot Optimum Pada Empat Variabel

Estimasi model untuk regresi semiparametrik *spline* pada IPKM Provinsi Jawa Timur adalah sebagai berikut.

$$\hat{y} = 0,73408 - 0,00246x_2 + 0,00006t_1 - 0,00091(t_1 - 2716)_+^1 + 0,0024(t_1 - 3382,86)_+^1 - 0,00156(t_1 - 3716,29)_+^1 - 0,00057t_3 + 0,00832(t_3 - 31,48)_+^1 - 0,02661(t_3 - 35,57)_+^1 + 0,01873(t_3 - 37,62)_+^1 + 0,00108t_4 - 0,02051(t_4 - 23,88)_+^1 + 0,061(t_4 - 30,41)_+^1 - 0,04178(t_4 - 33,68)_+^1$$

## H. Pengujian Asumsi Residual

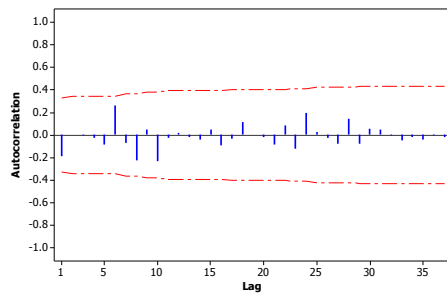
### 1. Pengujian Asumsi Identik

Pengujian asumsi identik pada residual dilakukan dengan menggunakan uji Glejser. Hasil uji Glejser akan menghasilkan nilai F sebesar 0,934 dan *P-value* sebesar 0,536. Dengan taraf

signifikansi ( $\alpha$ ) sebesar 5% maka didapatkan keputusan gagal tolak  $H_0$ , sehingga dapat disimpulkan bahwa tidak terjadi heteroskedastisitas atau dengan kata lain variansi antar residual sama.

## 2. Asumsi Independen

Salah satu cara untuk mengetahui ada atau tidaknya autokorelasi antar residual adalah dengan menggunakan Plot *Autocorrelation Function* (ACF). Gambar 2 menunjukkan bahwa tidak terlihat adanya nilai autokorelasi yang keluar batas interval konfidensi, sehingga dapat disimpulkan bahwa tidak terdapat autokorelasi antar residual.



Gambar 2. Plot ACF Residual

## 3. Pengujian Asumsi Distribusi Normal

Pengujian dilakukan dengan uji *Kolmogorov-Smirnov* dengan hasil nilai *Kolmogorov-Smirnov* sebesar 0,07. Nilai tersebut lebih kecil dibandingkan  $q_{(1-\alpha)}$  pada taraf signifikansi ( $\alpha$ ) sebesar 5%, yaitu 0,215, sehingga gagal tolak  $H_0$ . Hal ini menunjukkan bahwa residual telah memenuhi asumsi distribusi normal.

### I. Interpretasi Model Regresi Semiparametrik Spline

- Dengan mengasumsikan variabel prediktor selain  $x_2$  konstan, maka pengaruh variabel Angka Kematian Bayi terhadap IPKM di Provinsi Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = 0,73408 - 0,00246x_2$$

Apabila terjadi kenaikan Angka Kematian Bayi sebanyak satu persen, maka IPKM di Provinsi Jawa Timur akan turun sebesar 0,00246.

- Dengan mengasumsikan variabel prediktor selain  $t_1$  konstan, maka pengaruh variabel kepadatan penduduk terhadap IPKM di Provinsi Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = \begin{cases} 0,0000595t_1 & ; t_1 < 2716 \\ 2,46857 - 0,00085t_1 & ; 2716 \leq t_1 < 3382,86 \\ -5,64928 + 0,00155t_1 & ; 3382,86 \leq t_1 < 3716,29 \\ 0,15408 - 0,00001t_1 & ; t_1 \geq 3716,29 \end{cases}$$

Untuk kepadatan penduduk yang kurang dari 2716, apabila terjadi kenaikan pada kepadatan penduduk sebesar satu satuan, maka akan terjadi penurunan pada IPKM sebesar 0,0000595. Untuk kepadatan penduduk yang berada di antara 2716 hingga 3382,86, apabila terjadi kenaikan pada kepadatan penduduk sebesar satu satuan, maka akan terjadi kenaikan pada IPKM sebesar 0,00085. Selanjutnya, untuk kepadatan penduduk yang berada di antara 3382,86 hingga 3716,29, apabila terjadi kenaikan pada kepadatan penduduk sebesar satu satuan,

maka akan terjadi kenaikan pada IPKM sebesar 0,00155. Untuk kepadatan penduduk yang lebih dari 3716,29, apabila terjadi kenaikan pada kepadatan penduduk sebesar satu satuan, maka akan terjadi penurunan pada IPKM sebesar 0,00001.

- Dengan mengasumsikan variabel prediktor selain  $t_3$  konstan, maka pengaruh variabel persentase rumah tangga berperilaku hidup bersih dan sehat terhadap IPKM di Provinsi Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = \begin{cases} -0,00057t_3 & ; t_3 < 31,48 \\ -0,262 + 0,00776t_3 & ; 31,48 \leq t_3 < 35,57 \\ 0,68436 - 0,01885t_3 & ; 35,57 \leq t_3 < 37,62 \\ -0,02027 - 0,00012t_3 & ; t_3 \geq 37,62 \end{cases}$$

- Dengan mengasumsikan variabel prediktor selain  $t_4$  konstan, maka pengaruh variabel persentase rumah sehat terhadap IPKM di Provinsi Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = \begin{cases} 0,00108t_4 & ; t_4 < 23,88 \\ 0,48968 - 0,01942t_4 & ; 23,88 \leq t_4 < 30,41 \\ -1,36523 + 0,04157t_4 & ; 30,41 \leq t_4 < 33,68 \\ 0,04204 - 0,00021t_4 & ; t_4 \geq 33,68 \end{cases}$$

## V. KESIMPULAN DAN SARAN

Pada tahun 2013, IPKM terendah di Provinsi Jawa Timur yaitu Kabupaten Pamekasan yang disebabkan oleh kurangnya pengetahuan masyarakat tentang gizi bayi dan balita serta masih kurang sadar akan pentingnya perilaku hidup bersih dan sehat. Sedangkan Kota Madiun meraih IPKM tertinggi yang disebabkan oleh fasilitas kesehatan yang baik dari segi kualitas maupun kuantitas. Terdapat 13 kabupaten/kota di Provinsi Jawa Timur yang berada di bawah IPKM Indonesia. Model regresi semiparametrik *spline* terbaik dalam pemodelan IPKM di Provinsi Jawa Timur adalah dengan menggunakan tiga titik knot. Model ini memiliki empat variabel yang signifikan, yaitu Angka Kematian Bayi, kepadatan penduduk, persentase rumah tangga berperilaku hidup bersih dan sehat, dan persentase rumah sehat. Koefisien determinasi yang dihasilkan dari model ini adalah 92,99%. Hal tersebut menunjukkan bahwa model tersebut mampu menjelaskan keragaman IPKM Provinsi Jawa Timur sebesar 92,99%, sedangkan sisanya dijelaskan oleh variabel lain.

Saran yang dapat diberikan pemerintah adalah sebaiknya lebih memerhatikan faktor-faktor yang berpengaruh terhadap IPKM baik dari segi ekonomi, lingkungan, perilaku, dan kependudukan. Selain itu, dapat membuat program-program maupun kebijakan yang terkait dengan kesehatan sehingga dapat meningkatkan IPKM di Provinsi Jawa Timur.

## DAFTAR PUSTAKA

- BPS, Indeks Pembangunan Manusia 2013, Jakarta: Badan Pusat Statistik, 2013.
- Tim Penyusun IPKM, Indeks Pembangunan Kesehatan Masyarakat, Jakarta: Badan Penelitian dan Pengembangan Kesehatan, 2014.
- I. N. Budiantara, "Model Keluarga Spline Polinomial Truncated Dalam Regresi Semiparametrik", *Berkala Ilmiah MIPA Volume 15 No.3*,



- Yogyakarta: Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Gadjah Mada, 2005.
- [4] R. Riskiyanti, Analisis Regresi Multivariat Berdasarkan Faktor-Faktor yang Memengaruhi Derajat Kesehatan di Provinsi Jawa Timur, Surabaya: Institut Teknologi Sepuluh Nopember, 2010.
  - [5] A. Maully, Faktor-Faktor yang Memengaruhi Indeks Kesehatan Kabupaten dan Kota di Provinsi Jawa Timur, Surabaya: Institut Teknologi Sepuluh Nopember, 2014.
  - [6] D. Prasetyo, Pemodelan Data Kesehatan Kabupaten Banyuwangi dengan Regresi Terboboti Geografis, Bogor: Institut Pertanian Bogor, 2012.
  - [7] N. Diansuantari, Analisis Derajat Kesehatan Masyarakat Provinsi Bali dengan Menggunakan Metode Metode Multivariate Adaptive Regression Splines (MARS), Badung: Universitas Udayana, 2015.
  - [8] T. Kim, Y. Lee, and P. Newbold, Spurious Nonlinear Regressions in Econometrics, Nottingham NG7 2RD, UK: School of Economics, University of Nottingham, 2004.
  - [9] N. R. Drapper and H. Smith, Analisis Regresi Terapan, Diterjemahkan oleh: Bambang Sumantri, Jakarta: Gramedia Pustaka Utama, 1992.
  - [10] I. N. Budiantara, Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang, Pidato Pengukuhan Untuk Jabatan Guru Besar pada Jurusan Statistika, Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Surabaya: ITS Press, 2009.
  - [11] I. N. Budiantara, "Model Spline dengan Knots Optimal", *Jurnal Ilmu Dasar, FMIPA Universitas Jember*, Vol. 7, Hal. 77-85, 2006.
  - [12] D. Gujarati and D. Porter, Basic Econometrics, 5<sup>th</sup> Edition, New York: The McGraw-Hill Companies, Inc, 2009.
  - [13] W. Daniel, Statistika Nonparametrik Terapan, Diterjemahkan oleh: Alex Tri Kantjono W., Jakarta: PT Gramedia, 1989.